In the Specification

On page 37, line 29, before the semicolon, insert --who

states, "The linear sequence of amino acids in the protein has

the interesting property that it uniquely determines how [the]

chain will fold. The sequence, therefore, controls both the

protein's three-dimensional structure and ultimately its

function, which depends largely on the protein's shape....Knowing

what this three-dimensional structure looks like is crucial to

the process of designing drugs....Two spectroscopic techniques

help to determine the complex 3-D structures of proteins. The

first is X-ray crystallography....A second, and relatively new

technique, uses nuclear magnetic resonance (NMR) to detect which

atoms of the protein are brought close together by the folding

process. By combining these analytical data with building

models, we can arrive at a 3-D structure for the

protein....[Another idea] is to use the 3-D structures already

determined by X-ray crystallography and try to develop rules

based on how these proteins fold....To generate [a computer]

graphics display, we feed the coordinates or positions of the

atoms of the protein structure produced by X-ray, NMR or model

building into [a graphics] computer program. At this juncture,

the medicinal chemist and molecular modelers are at last ready to

begin the process of drug design by computer....The first and

most obvious [factor to consider] is that the small molecule must

2

have a geometrically complementary shape to the enzyme's active

site (the key must have an appropriate shape to fit the lock).

In addition, it must consist of appropriate atoms so that the

electronic characteristics of the molecule complements those of

the atoms making up the walls of the active site. In other

words, positively charged, polarised, groups of atoms in the

small molecule must fit negatively charged groups in the site,

not positively charged groups, which would repel each other. We

are just beginning to understand the rules governing the relative

importance of specific interactions. Finally, the compound must

be stable (stay around long enough to bind to the protein) and

relatively easy to synthesize to be a viable proposition."--

On page 37, line 30, before the semicolon, insert --who

state, "Rational [antiviral] drug design, defined as the directed

synthesis of new compounds based on an understanding of a

prototype drug/viral structural or functional protein interaction

at the atomic level, is only now becoming a viable alternative to

empirical screening....The most important ingredient in a drug

design program is the obvious need to predict the activity of a

molecule before the effort is invested to synthesize it. While a

certain amount of success can be had by viewing a ligand/protein

structure on a graphics terminal, the multiplicity of possible

conformations the ligand and protein can assume make it

exceedingly difficult to make accurate predictions with any

3

regularity. Recent developments in the use of thermodynamic calculations and molecular dynamics simulations permit consideration of previously impossible computational problems through the use of supercomputers and new computational approaches. The thermodynamic-cyclic perturbation approach to thermodynamic calculations is likely to be increasingly used in the future (McCammon, J.A. 1987, Computer-aided molecular design. *Science* 238:486-91.) Since the free energy change of binding of a drug to a protein or virus would be exceedingly difficult, if not impossible, to calculate, this method enables one to calculate the relative free energy change for the binding of two different compounds. These calculations will enable the researcher to predict whether a compound proposed for synthesis is likely to possess a greater binding affinity to the target. The second promising application of computational chemistry to drug design is molecular dynamics simulation (McCammon, J.A., Harvey, S.C. 1987, *Dynamics of Proteins and Nucleic Acids*. Cambridge, London, New York: Cambridge Univ. Press, pp. 234). The object of this exercise is to simulate the dynamics of the drug/protein interactions based on the dynamical trajectories of each atom. This type of analysis can help identify areas of the drug where considerable movement is occurring when bound to the target, suggesting that the conformation of the drug may need to be constrained to maximize activity. This analysis may also

4

suggest ways in which the drug is exerting its effect on the

protein."--.

On page 37, line 32, before the semicolon, insert --who

state, "**ChemStat** has been developed to allow the rapid

construction of a database containing molecular coordinates

together with computed and experimental parameters and biological

data....The database is stored on disk and comprises a number of

segments. Each segment is a molecular structure together with

associated information stored in 104 data fields. Some or all of

the segments may be read into the **Chem-X** data matrix where

information may be added, deleted or modified. Data may be

placed in the fields by calculating properties within **Chem-X** (or

interfaced programs) or by reading from an external file. Each

property to be calculated must first be defined in terms of a

ChemStat template file. This flexible method allows definitions

from the supplied 'library' of calculation types, user-supplied

Chem-X command files or external programs interfaced through

ChemLib. An automatic procedure is then invoked whereby each

property is calculated for each segment before the database is

updated. This will generally result in a data matrix containing

a vast amount of information which must be 'sifted' before it can

be used predictively. Key fields must be identified as being

associated with the required property. A novel method of data

reduction is employed by **ChemStat** to address this problem. One

field in the data matrix is specified as being the 'observed' field (generally the biological activity). By the automatic calculation of the appropriate correlation coefficients, those fields which correlate best with this field but are not highly inter-correlated can be listed and passed to the reduced data matrix. The data may be viewed graphically or in tabular form and fields and segments may be interactively included or excluded from the calculations. More sophisticated statistical techniques such as principal component analysis or pattern recognition are made available through an interface to some of the more popular commercial packages. Data from the reduced data matrix is passed to the program and new or modified data can be passed back to **Chem-X** via an intermediate file. Once the key fields have been identified, they can be used as predictor variables for multiple regression against the 'observed' field. **ChemStat** will determine the regression equation and calculate predicted values for the 'observed' variable from the values of the predictor variables for each segment....ChemStat, ChemLib and ChemQH are modules of the molecular modelling software Chem-X, developed and distributed by Chemical Design Limited, Oxford, England."--

On page 37, line 33, before the semicolon, insert --who state, "This paper examines the problem of automated structure generation at specified binding sites. The objective is to obtain molecular graphs that span the binding site and

6

incorporate predicted ligand points at their vertices. Three

approaches are considered: brute force techniques, subgraph

addition and spacer skeletons. A spacer skeletons is a device

for representing several distinct classes of compound within just

one structure. It is used to reduce the amount of searching that

must be done in the primary phase of structure generation....The

CDS [Cambridge Structural Database] can be searched to find all

the compounds that contain a certain molecular fragment and the

geometric parameters of this fragment in each compound can be

calculated. Certain geometric calculations can be performed

directly on the DATA file at the retrieval stage and the

redundant information on irrelevant parts of the molecule is

ignored. The GEOM program will compute angles, centroids,

vectors, planes and internal coordinates for a user-specified

fragment (Allen *et al.* 1979 [*Acta crystallogr.* B 35,2331-2339]).

This powerful facility can provide the user with a comprehensive

range of geometric parameters, which can then be analysed

statistically. This database is of great use in the building of

spacer skeletons designed to model many different

fragments....The results of each search through the database were

read into a geometric analysis program which retrieved the

coordinate data from the DATA file, plotted out the whole

molecule by using the PLUTO 78 package (Motherwell & Clegg 1978

[PLUTO78. Program for plotting molecular and crystal structures.

In *Cambridge crystallographic files*. University of Manchester

Regional Computer Centre.]) and calculated the least-squares

plane through the atoms in the fragments. The deviation of each

atom from the plane was computed together with the distance from

the atom to the centroid of the fragment. The results were

plotted as frequency histograms of (i) the root mean square

deviation and (ii) the distance of highly deviant atoms from the

centroid....The program MNET was written to provide a reliable

means of constructing spacer skeletons either round a starting

nucleus or from scratch."--/

On page 37, line 35, before the period, insert --who state,

"The most relevant precedent for base pairing in model systems is

that of Rich, [Kyogoku et al., *J. Am. Chem. Soc.* 1968, 90,4151-

4157,] using cyclohexyluracil binding to 9-ethyladenine in $CHCl_3$.

Systematic structural modifications in both components were made

in this study and revealed trends concerning steric effects and

acid-base effects. At the same time these systems were used to

examiner the kinetics of the base-pairing event. Aromatic

stacking interactions of simple bases were studied by Chan [Chan

et al., *Am. Chem. Soc.* 1964, 86, 4182. Schweitzer et al. *Ibid.*

1965, 87, 5241-5247. Iwahashi et al. *Ibid.* 1977, 99, 7761-7765.]

in aqueous solutions, whereas Tinoco et al. [Tinoco et al.,

Nature, New Biol. 1973, 246, 40-41] have developed a set of rules

for the sequence-specific hydrogen bonding and stacking

contributions of various base pairs to the stability of intact

nucleic acids. Our departure from previous model studies is made

possible by the construction of a new molecular shape which

permits both hydrogen bonding and aromatic stacking forces to act

simultaneously. The structural developments are a consequence of

the use of Kemp's [Kemp et al., J. Org. Chem. 1981, 46, 5140-

5143] triacid **3**, in which a U-shaped (diaxial) relationship

exists between any two carboxyl functions. Conversion of the

triacid to the imide acid chloride **4b** gives an acylating agent

that can be attached via amide or ester linkages to practically

any available aromatic surface. The resulting structure features

an aromatic plane which can be roughly parallel to that of the

atoms in the imide function; hydrogen bonding and stacking forces

converge from perpendicular directions to provide a

microenvironment complimentary to adenine derivatives. The same

structural features are also present in a model for thymine and

uracil recognition developed by Hamilton [Hamilton et al., J. Am.

Chem. Soc. 1987, 109, 5035-5036.]"--

On page 39, lines 27-28, please delete, without prejudice,

the words "are specifically incorporated by reference. Other"

and insert therein the words --and other--.

On page 40, delete lines 1-2.

9